[BoldFont=Fira Sans SemiBold]Fira Sans Book Fira Mono

popquizbox[1] colback=nipun-lightblue!10, colframe=nipun-blue, boxrule=2pt, arc=3pt, left=8pt, right=8pt, top=8pt, bottom=8pt, title= **Quick Quiz 1**, fonttitle=**, coltitle=nipun-white, colbacktitle=nipun-blue, enhanced, attach boxed title to top left=xshift=0pt, yshift=-2pt, boxed title style=arc=3pt, boxrule=0pt**

definitionbox[1] colback=nipun-green!8, colframe=nipun-green, boxrule=1.5pt, arc=2pt, left=6pt, right=6pt, top=6pt, bottom=6pt, title= **Definition: 1**, fonttitle=**, coltitle=nipun-white, colbacktitle=nipun-green**

examplebox[1] colback=nipun-orange!8, colframe=nipun-orange, boxrule=1.5pt, arc=2pt, left=6pt, right=6pt, top=6pt, bottom=6pt, title= **Example: 1**, fonttitle=**, coltitle=nipun-white, colbacktitle=nipun-orange**

keypointsbox colback=nipun-blue!8, colframe=nipun-blue, boxrule=1.5pt, arc=2pt, left=6pt, right=6pt, top=6pt, bottom=6pt, title= **Key Points**, fonttitle=**,

Consider this dataset

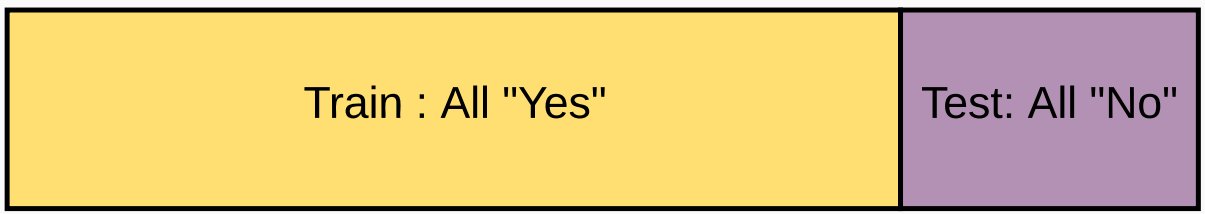
First 80 examples are of class "Yes"

Remaining 20 examples are of class "No".

Serial Number	**...**	**Class**
1		Yes
2		Yes
3		Yes
. .		. .
80		Yes
81		No
. .		. .
100		No

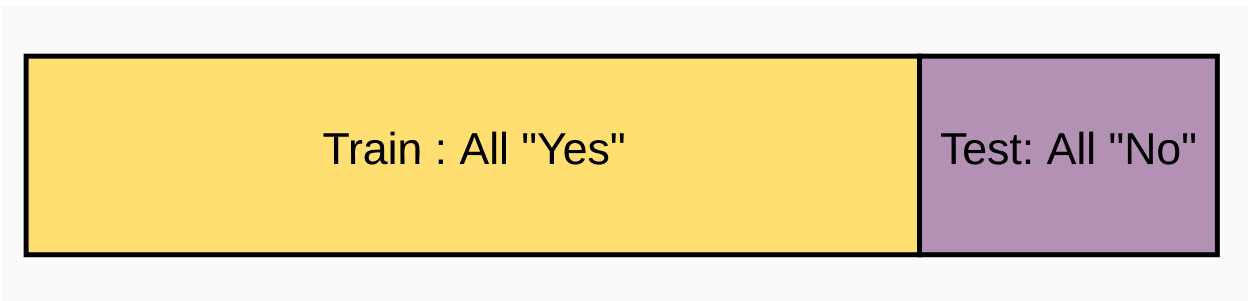
Consider this dataset

While using an 80-20 train-test split, we will get the distribution shown below



Consider this dataset

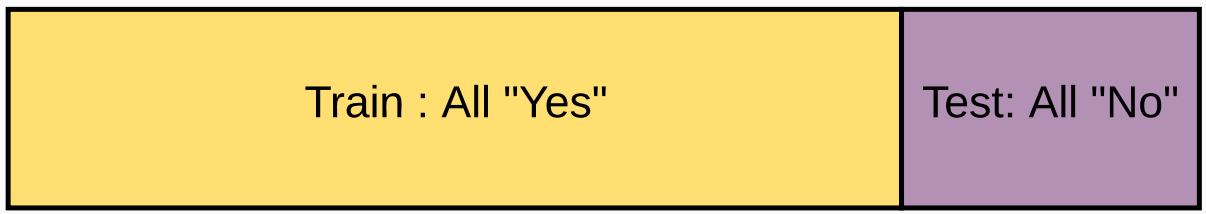
While using an 80-20 train-test split, we will get the distribution shown below



Will we learn anything useful in this scenario?

Consider this dataset

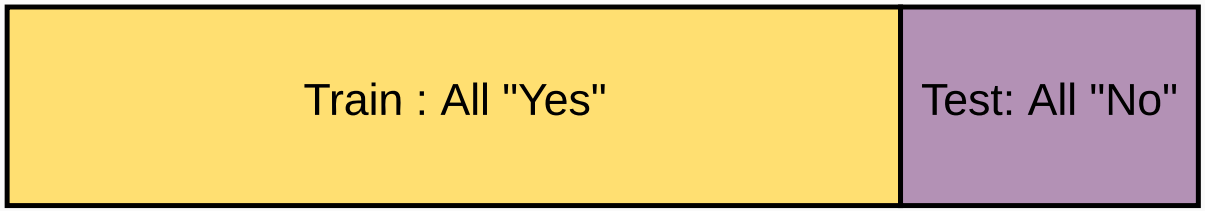
While using an 80-20 train-test split, we will get the distribution shown below



Will we learn anything useful in this scenario? *No* :(

Consider this dataset

While using an 80-20 train-test split, we will get the distribution shown below



Will we learn anything useful in this scenario? *No* :(

Solution: Shuffle before learning

Why shuffle for SGD?

We can fall into a loop!

SGD on point 1 : $\theta_0 + 0.2, \theta_1 - 0.2$
SGD on point 2 : $\theta_0 - 0.2, \theta_1 + 0.2$

Biased learning as point 2 follows point 1.